

	Motif 1	Motif 2	Motif 3	Motif 4
RV2837C_Nt	VGVVCHVPPDATTIGAGLALALVLDGGC	[35] VDLWTTVIPSVDRLGALG	[6] RELLYNDHEASND	[10] SADSTTMVAEILDAWGKPIDPRVAHTIAGLAATDGSSFRWAVS
S111253_Ssp	DLLIUCHOTEDFVVLGAAGVGLAKLHPGSR	[34] IRSLYITVDNOQGDRLGKA	[8] ROVALIDHRENSP	[11] AVGASTTILIVEKLORADISLSMVEASVMAIGIHWETGSJLTETQT
MCPA_Ng	IIVIEHHVRPDGCLGAQQGLFHLIKANF	[32] EALAIWWDANYXNRIELRE	[7] KAVLRIDHFPNED	[11] SYVACCEQIVEMATVAKWTTIPVVAATLTYGITYDSNRFLYSNT
YTQI_Bs	IIIEHRVREDPDAVGSCQQGLTEILRTRY	[30] GALVIVODTANOERIDQR	[4] AKLMKIDHFPNED	[13] SVSEMIYELLEGKEHGWKLNTKRAELIYAGIVGDTGRFLFPNT
AF2029_Af	LGIRTHDNPDPSMSSAYALREIAKQFD	[37] YDAFAIWDSGGPGVNNSIP	[3] DISIVIDHPAEK	[10] DVGATATILTEYIKELXKTPSKILATALLFFGKSETDEFKRNTR
MJ0977_Mj	NKLLIVTHHIDQDGLTSRAILQKLAERLN	[23] YDLIIFADLGSGQLKMKIE	[11] DKIIIIDHFOPEE	[18] GAEICGAGVSYLPKAINNDWIDLAKYAVLGAWGDIONIEGKLI
MJ1198_Mj	RPIIIRHHEDTGCGGTALEKAIPLII	[46] IPLIVLIDNGSTDDEDIPAI	[7] IEVIVIDHFPGE	[85] KGRTYDREYLEKIALCMDFEAFYLRFMDGKGIVDDILATNIKEF
HP1042_Hp	MQVFLSHIDPLGYACOLVSKOFFKNIQ	[27] EFLILNSDLNNLNEAEYL	[13] IOIQLIDHISGK	[19] IVYEFLKXHYAILEPKNTTWLEPLVEMVNSVDIWTQGYGFELS
RecJ_Hi	QKIVIVGDRDAGATSTALSVLRLQG	[31] VOLMLTVONGVSFSDGVAF	[5] IRLVLTIDHFPPE	[33] LAVRAKFRELGIFTAETOPNFTDLLDVALGTLADWVPLDQNNR
RecJ_Hp	TEIIIVVGDWDAAGVISSAIMAKFFESLN	[27] APLITIVDNGINAEAAFR	[5] YTLIITIDHFCHH	[27] LVAFILCYGHQLLGKEKSHEKSHSSSELLCLAGVATIADMPLTFNNR
RecJ_Ssp	EVKTIWGDFEPGITSTAVLWEGLQFFF	[32] TKLIVTCDGTGSTNLDEIVY	[5] MDVIVIDHFTLTD	[27] VAFKLVVEALYNQYPTVPRQPLEDILDIVAIGLIAADLVTLQGDCR
YBQ_Bs	ILIFGHONFDITICSAIAYADLKMKLG	[36] VNGVILVSHNERQQSIKDI	[3] QVLEVIDHRIAN	[12] PVGCATATILNKMYKENVKIEKEIAGLMLSATISDPSLILFKSPTC
ICRA_Sg	ILVPGHQNPDSDAIGSSYAFAYLARAY	[38] AEQVILTHNEFQQSVAIDI	[3] EYVGVNDHRRVAN	[12] PVGSASSIVYRMFKEHSVAVSKIEAGLMLSGLISDTLLKSPPT
MJ0608_Mj	RYWVGHKKNPDDSVCSAIIAYFLDCYP	[31] GKEIILVWHESEKQSFSDDL	[3] KLIAMIDHFKVGL	[19] IAELYFWDAILLIGGKKELKPDLAGLILSLSAISIDTVLFKSPPT
AF0756_Af	YVWVGHKKNPDDSVCSAIIAYFLWNWK	[46] GKXKVALVHSEKAOVTVDGI	[3] EVAVIDHFKIGD	[12] PVGCTATVILLFDKTGVEIPKDIAGILLSSLSLSDTIVIFKSATT
U60409_Lm	TVVQGNEGGDDSVVGCIYLAAMLFDKQP	[44] QIAHNHLVWIAALNASVULY	[14] RVVGVNDHFFDEQ	[11] LRTVGSACTLWTELRECGEDVVVCPTLITAPIVLDTVNFEPAQK
PPX1_Sc	TICVGNESADMSIASATTSYCYIYN	[62] ELSNSYLVWNNDPKLNKNY	[2] NWVGRIDBHDHFDLQ	[14] SCSSLVFNWYEKIQLQDREVVMNIAPLLMGALLIDITSNMRKVE
PRUNE_Dm	HLMAGNESQDLSAVSAVTLAFVYQAORH	[48] DVNVVILVWPHVSP LAPNVT	[1] NVEIIDHFPLED	[18] SVGSCATLVAVQRYLAEDQPRSTSVAQLLHATRVLDTINFAPAK
h-prune_1_Hs	HVWLGNEACDDSTVSALALAFYLAKT	[52] QLTLLVPHHILSKSDTAL	[2] AVAVIDHFPPIEP	[12] LVGSCATLVTERILQGAPEEILDRTAALLHGTTILDGVNMDLKG
GIS				
Specific motifs				
RV2837C_Nt	[107] TVNLA AVASGF GGCGHRLAAGYTTG	1648883	/	
S111253_Ssp	[123] DTDLTLQJLLEPYGCCHAQAAA VNL RDV	1653244	/	
MCPA_Ng	[102] GINVRDIAIKYGGCHNNASGAIITNK	1045875	/	
YTQI_Bs	[103] GPVINGLARKYNGGGHPLASGASIYSW	2293259	/	
AF2029_Af	[102] EVLRR AFGDVGSA CGGHAAGAQIPLG	2648507		Family 1
MJ0977_Mj	[254] PIKYASEKVKNGSGGGKFACGAYIPDN	2128614		
MJ1198_Mj	[109] QLMEEIPEASLDGCGGRECAGSLKFVEG	2128720		
HP1042_Hp	[137] CDVCELSQMCFNGGGHNRASGGKIDGF	2314198		
RecJ_Hi	[168] RIHSQHPNMILKFCCHAMAGLSIREE	1172895		
RecJ_Hp	[155] DALNGVSSLLGTYCCHEROACAGLSVEKN	2313437		
RecJ_Ssp	[165] ALLHSQRHMLGCGGHPFAGLSSLPLD	1652638		
YBQ_Bs	[31] DLSSKKTVEELLSIDAKEFTLG	[75] TALLKGVVSRKKQWVPLTD	586817	
ICRA_Sg	[31] NLASKSAEELIDIDAKTEELN	[75] HAFLAGAVSRKKQWVPLTE	1743856	
MJ0608_Mj	[32] VVGLKKEPELIMDFKNFDEN	[74] SVFLEGVMSRKKQWVPLLER	1591318	
AF0756_Af	[32] AVDDLTAMDIIKRDYKDFDMS	[75] SVWLDGVMSRKKQWVPLK	11498362	
U60409_Lm	[36] DVLA LSVPOLLRDYKOF SFK	[105] YSLSDFESI SRKK -LVP ALSE	1407725	
PPX1_Sc	[41] DIKGFSVS DILKDYKQENFQ	[97] MFQLQMVETRKQWVPLLEE	730369	
PRUNE_Dm	[37] DISKLTLTEVLRIDDMKV LQTD	[97] LRQHNVQATRKH -ILPIVKR	1079081	
h-prune_1_Hs	[34] DVSGLTT EQMLRKDQKTIYRQ	[94] YLQGNTQVSRKK -ILPLLQE	24308263	

Fig. 1